SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 6	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Statistically/Computationally Efficient Detection in Incompletely Characterized Non-Gaussian Noise via Parametric Modeling		5. TYPE OF REPORT & PERIOD COVERED August 1984 to July 1986
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Steven Kay and Debasis Sengupta		8. CONTRACT OR GRANT NUMBER(s) N00014-84-K-0527
9. PERFORMING ORGANIZATION NAME AND ADDRESS Electrical Engineering Department University of Rhode Island Kingston, Rhode Island 02881		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Virginia 22217		12. REPORT DATE August 1986
		13. NUMBER OF PAGES 35
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release. Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Non-Gaussian processes
Generalized likelihood ratio test
Rao efficient score test
Chi-square distribution
Autoregressive modeling
Maximum likelihood estimator
Signal to noise ratio

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A generalized likelihood ratio test is known to be able to reliably detect a signal known except for amplitude in incompletely characterized colored non-Gaussian noise, although it is computationally intensive. A Rao efficient score test is proposed as a computationally simpler alternative. The Rao test shares all the asymptotic properties of the generalized likelihood ratio test for large data records and small signal amplitudes. Its detection performance is asymptotically equivalent to that obtained for a similar detector designed

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Report Number 6

Statistically/Computationally Efficient Detection in Incompletely Characterized Colored Non-Gaussian Noise via Parametric Modeling

cop

STEVEN KAY AND DEBASIS SENGUPTA

Department of Electrical Engineering
University of Rhode Island University.
Kingston, Rhode Island 02881

August 1986

Prepared for

OFFICE OF NAVAL RESEARCH
Statistics and Probability Branch
Arlington, Virginia 22217
under Contract N00014-84-K-0527
S.M. Kay, Principal Investigator

Approved for public release; distribution unlimited

Abstract

A generalized likelihood ratio test is known to be able to reliably detect a signal known except for amplitude in incompletely characterized colored non-Gaussian noise, although it is computationally intensive. A Rao efficient score test is proposed as a computationally simpler alternative. The Rao test shares all the asymptotic properties of the generalized likelihood ratio test for large data records and small signal amplitudes. Its detection performance is asymptotically equivalent to that obtained for a similar detector designed with *a priori* knowledge of the unknown noise parameters. Computer simulations of the performance of the Rao detector support the theoretical results. A Rao detector built with the knowledge of the true form of the noise PDF is shown to significantly outperform a detector which assumes the noise to be Gaussian.

I. Introduction

Detection of a weak signal in noise is a problem of general interest, having been addressed by previous researchers. The problems studied range from detection of a completely known signal in known white Gaussian noise [Van Trees 1968] to detection of an unknown signal in unknown colored non-Gaussian noise. An attempt to solve the latter problem has been made recently by Kay and Sengupta [1986, 3]. It was assumed that the signal is known except for its amplitude which can be positive or negative. The problem is cast as testing of composite hypotheses and a generalized likelihood ratio test (GLRT) [Kendall and Stuart 1979] is employed to solve it a using parametric representation of the noise statistics. The GLRT is found to be particularly well-suited for this problem in that it has many attractive *asymptotic* properties such as consistency, unbiasedness and constant false alarm rate (CFAR). It is also asymptotically optimal in the sense that knowledge of the unknown noise parameters would not improve its performance provided the data record length is sufficiently large. However the GLRT has a serious disadvantage, namely, its computational complexity. It necessiates computing the maximum likelihood estimates (MLE) of all the unknown parameters under both the hypotheses. Computing the MLE of the signal amplitude is particularly difficult in the presence of other unknown parameters.

This paper proposes a Rao efficient score test [Rao 1948] as an alternative to the GLRT for the detection problem discussed above. The Rao test is shown to be equivalent to the GLRT when the signal amplitude is small. Consequently, it is also equivalent in performance to a *clairvoyant* Rao detector, which assumes the noise parameters to be known. It greatly reduces the computational complexity of the GLRT without sacrificing its optimality properties. A connection is established between the Rao detector and a locally optimal (LO) detector which assumes the polarity of the signal and the noise parameters to be known. Computer simulations support the theoretical predictions of the asymptotic performance of the Rao detector.

The paper is organized as follows. Section II introduces the Rao test as an asymptotic approximation to the GLRT. Section III presents two different noise models and defines the GLRT for each one. Section IV derives the corresponding Rao tests. Section V discusses the performance of the Rao test and compares it to the LO detector. Section VI derives the Rao detectors for the special cases of autoregressive (AR) noise with Gaussian and mixed-Gaussian distributions. Section VII reports the results of computer simulations and Section VIII sumarizes the main results.

II. GLRT and Rao Test

Let $\{u_1, u_2, \cdots, u_N\}$ be a set of independent and identically distributed (*i.i.d.*) random variables each having a probability density function (PDF) $f(u;\Theta)$ depending on the q-dimensional parameter vector $\Theta = [\theta_1 \ \theta_2 \ \cdots \ \theta_q]^T$. Consider the following composite hypothesis testing problem

$$\begin{aligned} \mathcal{H}_0 &: \Theta = [\mathbf{0}^T \ \Theta_s^T]^T \\ \mathcal{H}_1 &: \Theta = [\Theta_r^T \ \Theta_s^T]^T, \qquad \Theta_r \neq \mathbf{0} \end{aligned} \tag{1}$$

where Θ_r and Θ_s are r and s-dimensional parameter vectors, respectively, with $r+s=q$. $\mathbf{0}$ is an r-dimensional vector of 0's. Θ_s is assumed to be unknown. It is sometimes called the vector of *nuisance* parameters. The GLRT [Kendall and Stuart 1979] for testing (1) is to decide $\mathcal{H}_1$ if

$$\ell_G = \frac{\max\limits_{\Theta_r,\Theta_s} \prod\limits_{n=1}^{N} f(u_n;\Theta_r,\Theta_s)}{\max\limits_{\Theta_r} \prod\limits_{n=1}^{N} f(u_n;0,\Theta_s)} > \gamma \tag{2}$$

for some threshold γ. Let $\hat{\hat{\Theta}}_r$ and $\hat{\hat{\Theta}}_s$ be the MLE's of Θ_r and Θ_s under $\mathcal{H}_1$ and $\hat{\Theta}_s$ be the MLE of Θ_s under $\mathcal{H}_0$. Denoting the likelihood function by

$$\mathcal{L}(\Theta_r,\Theta_s) = \prod_{n=1}^{N} f(u_n;\Theta_r,\Theta_s),$$

the statistic (or *likelihood ratio*) for the GLRT can be written as

$$\ell_G = \frac{\mathcal{L}(\hat{\hat{\Theta}}_r, \hat{\hat{\Theta}}_s)}{\mathcal{L}(0, \hat{\Theta}_s)} \tag{3}$$

The quantity

$$V_i(\Theta_r, \Theta_s) = \frac{\partial}{\partial \theta_i} \ln \mathcal{L}(\Theta_r, \Theta_s)$$

is defined as the efficient score for the parameter θ_i and

$$\mathbf{V}(\Theta_r, \Theta_s) = [V_1(\Theta_r, \Theta_s)\ V_2(\Theta_r, \Theta_s)\ \cdots\ V_r(\Theta_r, \Theta_s)]^T$$

is the vector of efficient scores for the parameter vector Θ_r. It follows from the asymptotic properties of the MLE that under regularity conditions of the PDF [Kendall and Stuart 1979] $\mathbf{V}(\Theta_r, \Theta_s)$ is asymptotically of the form

$$\mathbf{V}(\Theta_r, \Theta_s) = \mathbf{I}(\Theta_r, \Theta_s)(\hat{\Theta} - \Theta) \tag{4}$$

where $\hat{\Theta}$ is the MLE of Θ and

$$[\mathbf{I}(\Theta_r, \Theta_s)]_{ij} = -E\left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln \mathcal{L}(\Theta_r, \Theta_s)\right] \qquad i, j = 1, 2, \cdots, q \tag{5}$$

which is the (i, j) element of the Fisher information matrix. Integration of (4) with respect to Θ results in [Cox and Hinkley 1974]

$$\mathcal{L}(\Theta_r, \Theta_s) = C \exp\left[-\frac{1}{2}(\hat{\Theta} - \Theta)^T \mathbf{I}(\Theta_r, \Theta_s)(\hat{\Theta} - \Theta)\right] \tag{6}$$

where C is a constant not dependent on Θ. The numerator and denominator of (3) correspond to the maximum of the right hand side of (6) under $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. The maximization with a few asymptotic arguments yield [Kendall and Stuart 1979] yeild

$$\ell_G = \frac{1}{\exp\left[-\frac{1}{2}\hat{\Theta}_r^T \mathbf{I}_{\Theta_r \Theta_r}(0, \hat{\Theta}_s)\hat{\Theta}_r\right]} \tag{7}$$

$\mathbf{I}_{\Theta_r\Theta_r}(\Theta_r, \Theta_s)$ is the $r \times r$ partition of $\mathbf{I}(\Theta_r, \Theta_s)$ corresponding to the parameter vector Θ_r or

$$[\mathbf{I}_{\Theta_r\Theta_r}(\Theta_r, \Theta_s)]_{ij} = -E\left[\frac{\partial^2}{\partial\theta_i\theta_j} \ln \mathcal{L}(\Theta_r, \Theta_s)\right] \qquad i, j = 1, 2, \cdots, r$$

Therefore for large data records (asymptotically) it follows from (7) that

$$2 \ln \ell_G = \hat{\Theta}_r^T \mathbf{I}_{\Theta_r\Theta_r}(\mathbf{0}, \hat{\Theta}_s)\hat{\Theta}_r \tag{8}$$

This result was proved rigorously by Wald [1943]. An alternative approximation of ℓ_G can be made by rewriting (6) the following way by using (4)

$$\mathcal{L}(\Theta_r, \Theta_s) = C \exp\left[-\frac{1}{2}\mathbf{V}^T(\Theta_r, \Theta_s)\mathbf{I}^{-1}(\Theta_r, \Theta_s)\mathbf{V}(\Theta_r, \Theta_s)\right] \tag{9}$$

The approximation of ℓ_G is

$$\ell_G = \exp\left[\frac{1}{2}\mathbf{V}_r^T(\mathbf{0}, \hat{\Theta}_s)\mathbf{J}(\mathbf{0}, \hat{\Theta}_s)\mathbf{V}_r(\mathbf{0}, \hat{\Theta}_s)\right] \tag{10}$$

where $\mathbf{V}_r(\Theta_r, \Theta_s)$ is the $r \times 1$ partition of $\mathbf{V}(\Theta_r, \Theta_s)$ corresponding to Θ_r,

$$\mathbf{V}_r(\Theta_r, \Theta_s) = [V_1(\Theta_r, \Theta_s)\ V_2(\Theta_r, \Theta_s)\ \cdots\ V_r(\Theta_r, \Theta_s)]^T$$

and $\mathbf{J}(\Theta_r, \Theta_s)$ is the $r \times r$ partition of $\mathbf{I}^{-1}(\Theta_r, \Theta_s)$

$$\mathbf{J}(\Theta_r, \Theta_s) = \left[\mathbf{I}_{\Theta_r\Theta_r}(\Theta_r, \Theta_s) - \mathbf{I}_{\Theta_r\Theta_s}(\Theta_r, \Theta_s)\mathbf{I}^{-1}_{\Theta_s\Theta_s}(\Theta_r, \Theta_s)\mathbf{I}^T_{\Theta_r\Theta_s}(\Theta_r, \Theta_s)\right]^{-1} \tag{11}$$

The terms in the brackets of are found by partitioning the Fisher information matrix for Θ

$$\mathbf{I}(\Theta) = \begin{pmatrix} \mathbf{I}_{\Theta_r\Theta_r}(\Theta_r, \Theta_s) & \mathbf{I}_{\Theta_r\Theta_s}(\Theta_r, \Theta_s) \\ \mathbf{I}_{\Theta_s\Theta_r}(\Theta_r, \Theta_s) & \mathbf{I}_{\Theta_s\Theta_s}(\Theta_r, \Theta_s) \end{pmatrix} \tag{12}$$

and the partitions are defined as

$$\mathbf{I}_{\Theta_r\Theta_r}(\Theta_r, \Theta_s) = E\left[\left(\frac{\partial \ln f}{\partial \Theta_r}\right)\left(\frac{\partial \ln f}{\partial \Theta_r}\right)^T\right] \qquad r \times r$$

$$\mathbf{I}_{\Theta_r\Theta_s}(\Theta_r, \Theta_s) = E\left[\left(\frac{\partial \ln f}{\partial \Theta_r}\right)\left(\frac{\partial \ln f}{\partial \Theta_s}\right)^T\right] \qquad r \times s$$

$$\mathbf{I}_{\Theta_s\Theta_r}(\Theta_r, \Theta_s) = \mathbf{I}^T_{\Theta_r\Theta_s}(\Theta_r, \Theta_s) \qquad s \times r$$

$$\mathbf{I}_{\Theta_s\Theta_s}(\Theta_r, \Theta_s) = E\left[\left(\frac{\partial \ln f}{\partial \Theta_s}\right)\left(\frac{\partial \ln f}{\partial \Theta_s}\right)^T\right] \qquad s \times s$$

This definition of the information matrices is equivalent to the previous definition (5). From (10) it follows that the following test statistic proposed by Rao [1948] is asymptotically equivalent to $2 \ln \ell_G$

$$\ell_R = \mathbf{V}_r^T(0, \hat{\Theta}_s)\mathbf{J}(0, \hat{\Theta}_s)\mathbf{V}_r(0, \hat{\Theta}_s) \tag{13}$$

Rao's test for the complex hypothesis testing problem (1) is to decide $\mathcal{H}_1$ if

$$\ell_R > \gamma' \tag{14}$$

where γ' is a suitable threshold and is equal to $2 \ln \gamma$ if the above asymptotic equivalence holds. The test is called Rao *efficient score* test since it uses the vector $\mathbf{V}_r$ of efficient score functions. Note that estimation of Θ_r *and* Θ_s under $\mathcal{H}_1$ is avoided by using the Rao test instead of the GLRT. This is an outcome of the approximation of a finite difference by a derivative as in (4) which holds only if the alternative hypothesis (H_1) tests for small departures from $\Theta_r = \mathbf{0}$. The computational simplicity of the Rao test makes it quite attractive for the composite hypothesis testing problem described above. A more rigorous derivation is available in [Rao 1973].

The statistics of ℓ_R are difficult to obtain in general. For large data records (asymptotically) it may be shown that both $2 \ln \ell_G$ and ℓ_R are distributed in the following manner [Rao 1973].

$$\ell_R \sim \chi_r^2 \qquad \text{under } \mathcal{H}_0 \tag{15a}$$

$$\ell_R \sim \chi'^2(r, \lambda) \qquad \text{under } \mathcal{H}_1 \tag{15b}$$

Here χ_r^2 represents a chi-square distribution with r degrees of freedom and $\chi'^2(r, \lambda)$ represents a noncentral chi-square distribution with r degrees of freedom and noncentrality parameter λ. Note that $\chi'^2(r, 0) = \chi_r^2$ or the distribution under $\mathcal{H}_0$ is a special case of the distribution under $\mathcal{H}_1$ and occurs when $\lambda = 0$. The noncentrality parameter λ, which is a measure of the discrimination between two hypotheses, is given by

$$\lambda = \Theta_r^T \left[\mathbf{I}_{\Theta_r\Theta_r}(0, \Theta_s) - \mathbf{I}_{\Theta_r\Theta_s}(0, \Theta_s)\mathbf{I}_{\Theta_r\Theta_s}^{-1}(0, \Theta_s)\mathbf{I}_{\Theta_r\Theta_s}^T(0, \Theta_s) \right] \Theta_r \tag{16}$$

All the terms in the brackets of (16) are partitions of $\mathbf{I}(\Theta_r, \Theta_s)$ as given by (12).

III. The Detection Problem and the GLRT Solution

IIIA. The General Linear Model

Consider the following detection problem.

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} = \mathbf{W}\mathbf{u} \\ \mathcal{H}_1 &: \mathbf{y} = \mathbf{W}\mathbf{u} + \mu \mathbf{s} \end{aligned} \tag{17}$$

where $\mathbf{s} = [s_1\ s_2\ \cdots\ s_N]^T$ is a vector of known signal amplitudes, $\mathbf{u} = [u_1\ u_2\ \cdots\ u_N]^T$ is a vector of *i.i.d.* noise with a symmetric PDF, μ is an unknown scalar (either positive or negative) and $\mathbf{W}$ is an invertible $(N \times N)$ matrix whose elements are functions of a set of unknown parameters $\Psi = [\psi_1\ \psi_2\ \cdots\ \psi_M]$.

$$[\mathbf{W}]_{ij} = \mathrm{w}_{ij}(\Psi)$$

Since $u_n,\ \ n = 1, 2, \cdots N$ are *i.i.d.*, the PDF of $\mathbf{u}$ can be expressed as

$$\mathbf{f}(\mathbf{u}; \Phi) = \prod_{n=1}^{N} f(u_n; \Phi) \tag{18}$$

where $f(u_n; \Phi)$ is the marginal PDF of each u_n dependent on the unknown parameter vector Φ. f is assumed to be an even PDF, *i.e.*, $f(-u) = f(u)$.

The linear model of (17) is capable of representing a large class of correlation patterns of the background noise. The assumption of a known PDF with unknown parameters Φ adds flexibility to the model while still maintaining the parametric form. A detector based on this model would be insensitive to a change of polarity of the signal since μ can be positive or negative. (17) is written as the hypothesis testing problem

$$\mathcal{H}_0 : \Theta^T = [\mathbf{0}^T\ \Theta_s^T] \tag{19a}$$

$$\mathcal{H}_1 : \Theta^T = [\mathbf{0}^T\ \Theta_s^T] \qquad\qquad \Theta_r \neq \mathbf{0} \tag{19b}$$

where

$$\begin{aligned} \Theta_r &= \mu \qquad \text{(a scalar)} \\ \Theta_s &= [\Psi^T \ \Phi^T]^T \qquad \text{(vector of } \textit{nuisance} \text{ parameters)} \end{aligned} \tag{20}$$

The vector $\mathbf{y}$ is a linear function of the vector of *i.i.d.* random variables $\mathbf{u}$ under either hypothesis. Using this fact the joint PDF of $\mathbf{y}$ is found to be

$$\mathbf{f}(\mathbf{y}; \Psi, \Phi) = \frac{1}{|\det(\mathbf{W})|} \prod_{n=1}^{N} \left(f(u_n; \Phi) \Big|_{u_n = \sum_{j=1}^{N} \omega_{nj}(\Psi) y_j} \right) \qquad \textit{under } \mathcal{H}_0 \tag{21a}$$

$$\mathbf{f}(\mathbf{y}; \Psi, \Phi) = \frac{1}{|\det(\mathbf{W})|} \prod_{n=1}^{N} \left(f(u_n; \Phi) \Big|_{u_n = \sum_{j=1}^{N} \omega_{nj}(\Psi)(y_j - \mu s_j)} \right) \textit{under } \mathcal{H}_1 \tag{21b}$$

where $\omega_{nj}(\Psi)$ are elments of $\mathbf{W}^{-1}$, which are known functions of Ψ.

$$\omega_{nj}(\Psi) = [\mathbf{W}^{-1}]_{nj}$$

The GLRT for testing (17) has been shown to be [Kay and Sengupta 1986, 3] equivalent to deciding $\mathcal{H}_1$ if

$$\ell_G = \frac{\prod_{n=1}^{N} f\left(\sum_{j=1}^{N} \omega_{nj}(\hat{\hat{\Psi}})(y_j - \hat{\mu} s_j); \hat{\hat{\Phi}} \right)}{\prod_{n=1}^{N} f\left(\sum_{j=1}^{N} \omega_{nj}(\hat{\Psi}) y_j; \hat{\Phi} \right)} > \gamma \tag{22}$$

$\hat{\hat{\Psi}}$, $\hat{\mu}$ and $\hat{\hat{\Phi}}$ are joint MLE's of Ψ, μ and Φ, respectively, under $\mathcal{H}_1$. $\hat{\Psi}$ and $\hat{\Phi}$ are joint MLE's of Ψ and Φ under $\mathcal{H}_0$.

IIIB. The AR Noise Model

The detection problem for AR noise is

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} = \mathbf{x} \\ \mathcal{H}_1 &: \mathbf{y} = \mathbf{x} + \mu \mathbf{s} \end{aligned} \tag{23}$$

with

$$\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_N]^T$$

It is assumed that the sequence $\{x_1, x_2, \cdots, x_N\}$ is the output of a p-th order all-pole filter excited by white driving noise or

$$x_n = -\sum_{j=1}^{p} a_j x_{n-j} + u_n, \qquad n = 1, 2, \cdots, N$$

Alternately,

$$u_n = \sum_{j=0}^{p} a_j x_{n-j}, \qquad n = 1, 2, \cdots, N$$

assuming $a_0 = 1$. If the samples $y_{-p+1}, y_{-p+2}, \cdots, y_0$ are assumed to be 0, then (23) can be shown to be aspecial case of (7) where $\mathbf{W}^{-1}$ is a lower triangular Toeplitz matrix [Kay and Sengupta 1986, 3]. The results of the previous section can then be used to determine the GLRT. Alternatively, the GLRT can be derived on the basis of the *conditional* likelihood function, assuming the data records to be large. The conditional likelihood of $y_{p+1}, y_{p+2}, \cdots, y_N$ given $y_1, y_2, \cdots, y_p$ is

$$\begin{aligned} & \mathbf{f}(y_{p+1}, y_{p+2}, \cdots, y_N | y_1, y_2, \cdots, y_p) \\ & = \prod_{n=p+1}^{N} f\left(\sum_{j=0}^{p} a_j y_{n-j}; \Phi\right) \qquad \text{under } \mathcal{H}_0 \qquad (24a) \\ & = \prod_{n=p+1}^{N} f\left(\sum_{j=0}^{p} a_j (y_{n-j} - \mu s_{n-j}); \Phi\right) \qquad \text{under } \mathcal{H}_1 \qquad (24b) \end{aligned}$$

The GLRT is then given by

$$\ell_G = \left[\frac{\prod_{n=p+1}^{N} f\left(\sum_{j=0}^{p} \hat{\hat{a}}_j (y_{n-j} - \hat{\hat{\mu}} s_{n-j}); \hat{\hat{\Phi}}\right)}{\prod_{n=p+1}^{N} f\left(\sum_{j=0}^{p} \hat{a}_j y_{n-j}; \hat{\Phi}\right)} \right] \qquad (25)$$

where hat's indicate MLE's under $\mathcal{H}_0$ and double hat's indicate MLE's under $\mathcal{H}_1$, respectively. $\hat{\hat{a}}_0$ and $\hat{a}_0$ are defined to be unity.

Computation of (22) or (25) involves evaluation of the MLE's of all the parameters under H_1. The Rao test, as indicated in the previous section, is able to avoid this computation. The following section illustrates how this is accomplished.

IV. Rao Test for the Detection Problem

IVA. The General Linear Model

For the detection problem of (17) Θ_r and Θ_s are given by (20). In this case ℓ_R defined in (13) can be computed from the likelihood function or PDF of $\mathbf{y}$ (see (21)). It is observed that

$$\begin{aligned}
\mathbf{V}_r(0,\Theta_s) &= \left[\frac{\partial \ln \mathbf{f}(\mathbf{y};\Theta_r,\Theta_s)}{\partial \Theta_r}\right]\Bigg|_{\Theta_r=0} \\
&= \frac{\partial}{\partial \mu}\ln \mathbf{f}(\mathbf{y};\mu,\Theta_s)\Bigg|_{\mu=0} \\
&= \frac{\partial}{\partial \mu}\ln\left[\prod_{n=1}^{N} f\left(\sum_{j=1}^{N}\omega_{nj}(\Psi)(y_j-\mu s_j);\Phi\right)\right]\Bigg|_{\mu=0} \\
&= \frac{\partial}{\partial \mu}\ln\left[\prod_{n=1}^{N} f(u_n;\Phi)\right]\Bigg|_{\mu=0}
\end{aligned}$$

since from (17) u_n can be written as

$$u_n = \sum_{j=1}^{N}\omega_{nj}(y_j-\mu s_j) \tag{26}$$

under $\mathcal{H}_1$. Therefore

$$\begin{aligned}
\mathbf{V}_r(0,\Theta_s) &= \sum_{n=1}^{N}\left[\frac{\partial}{\partial \mu}\ln f(u_n;\Phi)\right]\Bigg|_{\mu=0} \\
&= \sum_{n=1}^{N}\left[\left(\frac{\partial u_n}{\partial \mu}\right)\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\right]\Bigg|_{\mu=0} \\
&= \sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\Psi)s_j\right)\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\Bigg|_{\mu=0} \\
&= \sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\Psi)s_j\right)\left[\frac{f'(u_n;\Phi)}{f(u_n;\Phi)}\right]\Bigg|_{\mu=0}
\end{aligned} \tag{27}$$

It has been proved for the detection problem considered here (and specifically if $f(u)$ is an even function) that [Sengupta 1986] under certain regularity conditions

$$\mathbf{I}_{\Theta_r\Theta_s}(\Theta_r,\Theta_s) = \mathbf{I}_{\mu\Theta_s}(\mu,\Theta_s) = \mathbf{0}$$

Hence from (11) it follows that

$$\begin{aligned}
\mathbf{J}^{-1}(\Theta_r,\Theta_s) &= I_{\mu\mu}(\mu,\Psi,\Phi)\\
&= \sum_{n=1}^{N} E\left[\left\{\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\left(\frac{\partial u_n}{\partial \mu}\right)\right\}^2\right]\\
&= \sum_{n=1}^{N} E\left[\left(-\sum_{j=1}^{N}\omega_{nj}s_j\right)^2\left(\frac{\partial \ln f}{\partial u_n}\right)^2\right]
\end{aligned}$$

i.e.,

$$\begin{aligned}
\mathbf{J}^{-1}(0,\Theta_s) &= I_{\mu\mu}(0,\Psi,\Phi)\\
&= \left[\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}s_j\right)^2\right] \mathcal{I}_f(\Phi) \qquad (28)
\end{aligned}$$

where

$$\mathcal{I}_f(\Phi) = E\left[\left(\frac{\partial \ln f}{\partial u_n}\right)^2\right]\Bigg|_{\mu=0} \qquad (29)$$

Substituting (27) and (28) in (13)

$$\ell_R = \frac{\left[\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})s_j\right)\left[\frac{f'\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}{f\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}\right]\right]^2}{\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})s_j\right)^2 \mathcal{I}_f(\hat{\Phi})} \qquad (30)$$

If Φ is a scalar, $\mathcal{I}_f(\hat{\Phi})$ can be stored in a table as a function of the parameter. But if Φ is a vector consisting of multiple parameters, it is a difficult task to evaluate it by integration for each value of $\hat{\Phi}$. On the other hand it is possible to use the asymptotic

equivalence [Kay 1985]

$$\mathcal{I}_f(\hat{\Phi}) = E\left[\left(\frac{f'}{f}\right)^2\right]$$
$$\approx \frac{1}{N}\sum_{n=1}^{N}\left[\frac{f'\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}{f\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}\right]^2$$

Substituting in (30)

$$\ell_R = \frac{\left[\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})s_j\right)\left[\frac{f'\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}{f\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}\right]\right]^2}{\frac{1}{N}\left[\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})s_j\right)^2\right]\sum_{n=1}^{N}\left[\frac{f'\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}{f\left(\sum_{j=1}^{N}\omega_{nj}(\hat{\Psi})y_j;\hat{\Phi}\right)}\right]^2} \qquad (31)$$

In the case of *white* noise

$$\omega_{nj} = \begin{cases} 1, & n = j,\ n, j = 1, 2, \cdots, N \\ 0, & n \neq j,\ n, j = 1, 2, \cdots, N \end{cases}$$

and (31) reduces to

$$\ell_R = \frac{\left[\sum_{n=1}^{N} s_n \left[\frac{f'(y_n;\hat{\Phi})}{f(y_n;\hat{\Phi})}\right]\right]^2}{\frac{1}{N}\left(\sum_{n=1}^{N} s_n^2\right)\sum_{n=1}^{N}\left[\frac{f'(y_n;\hat{\Phi})}{f(y_n;\hat{\Phi})}\right]^2}$$

as obtained by Kay [1985].

IVB. The AR Noise Model

The parameter vector in the AR noise case is

$$\Theta_r = \mu \qquad \text{(a scalar)}$$

$$\Theta_s = [\mathbf{a}^T \ \Phi^T]^T \qquad \text{(vector of } \textit{nuisance} \text{ parameters)}$$

In this case it is easier to use the conditional likelihood functions given by (24). The vector of efficient score functions is found to be

$$\begin{aligned}
\mathbf{V}_r(\mathbf{0},\Theta_s) &= \frac{\partial}{\partial\mu}\ln\left[\prod_{n=p+1}^{N} f\left(\sum_{j=0}^{p} a_j(y_{n-j}-\mu s_{n-j});\Phi\right)\right]\Bigg|_{\mu=0} \\
&= \frac{\partial}{\partial\mu}\ln\left[\prod_{n=p+1}^{N} f(u_n;\Phi)\right]\Bigg|_{\mu=0}
\end{aligned}$$

where u_n can be written from (23) as

$$u_n = \sum_{j=0}^{p} a_j(y_{n-j}-\mu s_{n-j}) \tag{32}$$

under $\mathcal{H}_1$. Hence

$$\begin{aligned}
\mathbf{V}_r(\mathbf{0},\Theta_s) &= \sum_{n=p+1}^{N}\left[\frac{\partial}{\partial\mu}\ln f(u_n;\Phi)\right]\Bigg|_{\mu=0} \\
&= \sum_{n=p+1}^{N}\left[\left(\frac{\partial u_n}{\partial\mu}\right)\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\right]\Bigg|_{\mu=0} \\
&= \sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p} a_j s_{n-j}\right)\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\Bigg|_{\mu=0} \\
&= \sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p} a_j s_{n-j}\right)\left[\frac{f'(u_n;\Phi)}{f(u_n;\Phi)}\right]\Bigg|_{\mu=0}
\end{aligned} \tag{33}$$

For the case of AR noise it has been proved that for f(u) an even function and under certain regularity conditions [Sengupta 1986]

$$\mathbf{I}_{\Theta_r\Theta_s}(\Theta_r,\Theta_s) = \mathbf{I}_{\mu\Theta_s}(\mu,\Theta_s) = \mathbf{0}$$

Hence from (11)

$$\begin{aligned}
\mathbf{J}^{-1}(\Theta_r, \Theta_s) &= I_{\mu\mu}(\mu, \mathrm{a}, \Phi) \\
&= \sum_{n=p+1}^{N} E\left[\left\{\left(\frac{\partial}{\partial u_n}\ln f(u_n;\Phi)\right)\left(\frac{\partial u_n}{\partial \mu}\right)\right\}^2\right] \\
&= \sum_{n=p+1}^{N} E\left[\left(-\sum_{j=0}^{p} a_j s_{n-j}\right)^2 \left(\frac{\partial \ln f}{\partial u_n}\right)^2\right]
\end{aligned}$$

i.e.,

$$\begin{aligned}
\mathbf{J}^{-1}(0, \Theta_s) &= I_{\mu\mu}(0, \mathrm{a}, \Phi) \\
&= \left[\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p} a_j s_{n-j}\right)^2\right] I_f(\Phi)
\end{aligned} \tag{34}$$

where $I_f(\Phi)$ is as defined in (29). Substituting (33) and (34) in (13)

$$\ell_R = \frac{\left[\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}\hat{a}_j s_{n-j}\right)\left[\frac{f'\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}{f\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}\right]\right]^2}{\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}\hat{a}_j s_{n-j}\right)^2 I_f(\hat{\Phi})} \tag{35}$$

As before, evaluation of the integration involved in $I_f(\Phi)$ can be avoided by the asymptotic equivalence

$$\begin{aligned}
I_f(\hat{\Phi}) &= E\left[\left(\frac{f'}{f}\right)^2\right] \\
&\approx \frac{1}{N-p}\sum_{n=p+1}^{N}\left[\frac{f'\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}{f\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}\right]^2
\end{aligned}$$

Substituting in (35)

$$\ell_R = \frac{\left[\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}\hat{a}_j s_{n-j}\right)\left[\frac{f'\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}{f\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}\right]\right]^2}{\frac{1}{N-p}\left[\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}\hat{a}_j s_{n-j}\right)^2\right]\sum_{n=p+1}^{N}\left[\frac{f'\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}{f\left(\sum_{j=0}^{p}\hat{a}_j y_{n-j};\hat{\Phi}\right)}\right]^2} \tag{36}$$

Either (35) or (36) can be convenient for use depending on the number of unknown PDF parameters.

V. Asymptotic Performance of the Rao Detector

Asymptotic distributions of ℓ_R under $\mathcal{H}_0$ and $\mathcal{H}_1$ are given by (15a) and (15b), respectively. For the general linear model $\Theta_r = \mu$ and $\Theta_s = [\Psi^T\ \Phi^T]^T$ while for the AR noise model $\Theta_r = \mu$ and $\Theta_s = [\mathrm{a}^T\ \Phi^T]^T$. Therefore in either case the noncentrality parameter is

$$\begin{aligned}\lambda &= \mu^2[I_{\mu\mu}(0,\Theta_s) - \mathbf{I}_{\mu\Theta_s}(0,\Theta_s)\mathbf{I}^{-1}_{\Theta_s\Theta_s}(0,\Theta_s)\mathbf{I}^T_{\mu\Theta_s}(0,\Theta_s)]\\ &= \mu^2 I_{\mu\mu}(0,\Theta_s)\end{aligned} \tag{37}$$

since $\mathbf{I}_{\mu\Theta_s}(0,\Theta_s) = 0$, as indicated before. The probability of false alarm is

$$P_{FA} = \mathcal{P}\{\ell_R > \gamma' | \mathcal{H}_0\} \tag{38a}$$

and the probability of detection is

$$P_D = \mathcal{P}\{\ell_R > \gamma' | \mathcal{H}_1\} \tag{38b}$$

Both the probabilities can be calculated from the tables of noncentral and central chi-square distributions, respectively. γ' can be set to produce a given false alarm rate and P_D can be calculated accordingly.

It is known that the GLRT for the detection problems considered here is asymptotically optimal in the sense that its performance is equivalent to that of a *clairvoyant* GLRT built with perfect knowledge of Θ_s provided the data record is large [Kay and Sengupta 1986, 3]. This result applies to the Rao test as well, since it is asymptotically equivalent to the GLRT. The clairvoyant Rao test assumes that Ψ and Φ are known or that from (30)

$$\ell_{RC} = \frac{\left[\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\Psi)s_j\right)\left[\frac{f'\left(\sum_{j=1}^{N}\omega_{nj}(\Psi)y_j;\Phi\right)}{f\left(\sum_{j=1}^{N}\omega_{nj}(\Psi)y_j;\Phi\right)}\right]\right]^2}{\sum_{n=1}^{N}\left(-\sum_{j=1}^{N}\omega_{nj}(\Psi)s_j\right)^2 I_f(\Phi)} \tag{39}$$

for the general linear model and in the AR noise case

$$\ell_{RC} = \frac{\left[\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}a_j s_{n-j}\right)\left[\frac{f'\left(\sum_{j=0}^{p}a_j y_{n-j};\Phi\right)}{f\left(\sum_{j=0}^{p}a_j y_{n-j};\Phi\right)}\right]\right]^2}{\sum_{n=p+1}^{N}\left(-\sum_{j=0}^{p}a_j s_{n-j}\right)^2 I_f(\Phi)} \tag{40}$$

Note that all the parameters are assumed to be known. Asymptotically ℓ_{RC} is distributed as

$$\ell_{RC} \sim \chi_r^2 \qquad \text{under } \mathcal{H}_0 \tag{41a}$$

$$\ell_{RC} \sim \chi'^2(r, \lambda_c) \qquad \text{under } \mathcal{H}_1 \tag{41b}$$

where

$$\lambda_c = \Theta_r^T \mathbf{I}_{\Theta_r\Theta_r}(0, \Theta_s)\Theta_r$$

As indicated before, λ and λ_c are asymptotically equivalent for both the models considered in this paper. Therefore

$$\lambda = \lambda_c = \mu^2 I_{\mu\mu}(0, \Theta_s) \tag{42}$$

For the general linear model substitution of (28) in (42) produces

$$\lambda = \left[\frac{\mu^2}{\sigma^2} \sum_{n=1}^{N} \left(-\sum_{j=1}^{N} \omega_{nj} s_j \right)^2 \right] \sigma^2 I_f(\Phi) \tag{43}$$

while in the AR case (35) can be substituted in (42) to obtain

$$\lambda = \left[\frac{\mu^2}{\sigma^2} \sum_{n=p+1}^{N} \left(-\sum_{j=0}^{p} a_j s_{n-j} \right)^2 \right] \sigma^2 I_f(\Phi) \tag{44}$$

σ^2 is the variance of u_n and has been added to facilitate the interpretation of λ. The noncentrality parameter is found to dependent on the noise PDF only through the quantity $\sigma^2 I_f$, assuming all the PDF's under consideration have the same variance. It is known that this quantity is greater than unity for all non-Gaussian PDF's [Kay and Sengupta 1986, 1]. For a Gaussian PDF $\sigma^2 I_f = 1$. Therefore all other parameters remaining the same, a non-Gaussian noise background will produce a larger value of λ than a Gaussian background, leading to a larger probability of detection. It can be shown that for non-Gaussian and Gaussian noise backgrounds with identical power spectral densities (PSD), $\sigma^2 I_f$ is the ratio of the SNR necessary in the Gaussian case to the SNR necessary in the Gaussian case, in order to produce a given probability of detection [Kay and Sengupta 1986, 3]. Section VII reports the results of computer simulations which illustrates this.

It is of interest to notice the connection of the Rao detector with a *locally optimum* (LO) detectors [Middleton 1966], [Czarnecki and Thomas 1984]. A LO detector for the

detection of a known signal with unknown but *positive* amplitude in *known* colored noise maximizes the probability of detection in the neighborhood of $\mu = 0$ and is represented by the statistic

$$T_{LO} = \lim_{\mu \to 0} \frac{\partial}{\partial \mu} \ln \mathbf{f}(\mathbf{y}; \mu, \Theta_s) \tag{45}$$

with usual notations. The test decides $\mathcal{H}_1$ if

$$T_{LO} > \gamma_{LO}$$

where γ_{LO} is an appropriate threshold. An equivalent test is to decide $\mathcal{H}_1$ if

$$\frac{T_{LO}}{\sqrt{I_{\mu\mu}(0, \Theta_s)}} = T'_{LO} > \gamma'_{LO} = \gamma / \sqrt{I_{\mu\mu}(0, \Theta_s)} \tag{46}$$

Since the performance of the Rao detector is asymptotically equivalent to a corresponding clairvoyant detector, it is possible to do a comparison on the basis of *known* nuisance parameters. From (13), ℓ_{RC} is given by

$$\begin{aligned} \ell_{RC} &= \left[\mathbf{V}_r^T(0, \Theta_s)\mathbf{J}(0, \Theta_s)\mathbf{V}_r(0, \Theta_s)\right] \\ &= \frac{\left[\frac{\partial}{\partial \mu} \ln \mathbf{f}(\mathbf{y}; 0, \Theta_s)\right]^2}{I_{\mu\mu}(0, \Theta_s)} \end{aligned}$$

which is observed to be the square of T'_{LO}. T_{LO} is Gaussian for large data records [Middleton 1966] while $\mathcal{L}_{RC}$ is chi-square distributed with a single degree of freedom. Therefore for a given probability of false alarm the locally optimal detector will have a slightly larger probability of detecting small amplitudes of signal. This makes intuitive sense, because (46) is a one-sided test while GLRT and the Rao detectors allow for positive *and* negative amplitudes of the signal. This is an example of trading off performance in order to make the test two-sided. This comparison applies to the case of known nuisance parameters only. In the case of unknown nuisance parameters the LO detector is not defined, while the Rao detector exists and achieves the performance of a clairvoyant detector asymptotically. Furthermore, the LO detectors for positive

and negative values of μ are different (*i.e.*, the inequality is reversed for negative amplitudes) and hence they are impractical in situations where the polarity of the signal can change.

VI. Rao Test for Gaussian and Mixed-Gaussian Noise

VIA. The Gaussian Case

The Rao detector is now derived for the special case of the zero mean Gaussian noise. Attention is restricted to AR noise only because of the abundance of available results [Bowyer 1979], [Kay 1983]. In this case $\Phi = \sigma^2$, a scalar, and

$$f(u_n;\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{u_n^2}{2\sigma^2}\right)}$$

such that

$$\frac{f'(u_n;\sigma^2)}{f(u_n;\sigma^2)} = -\frac{u_n}{\sigma^2} \tag{47}$$

Substituting (47) in (35) and using the fact that $I_f(\sigma^2) = 1/\sigma^2$ for a Gaussian PDF,

$$\ell_R = \frac{\left[\sum_{n=p+1}^{N}\left(s_n + \sum_{j=1}^{p}\hat{a}_j s_{n-j}\right)\left(y_n + \sum_{j=1}^{p}\hat{a}_j y_{n-j}\right)\right]^2}{\hat{\sigma}^2 \sum_{n=p+1}^{N}\left(s_n + \sum_{j=1}^{p}\hat{a}_j s_{n-j}\right)^2} \tag{48}$$

where

$$\hat{\sigma}^2 = \frac{1}{N-p}\sum_{n=p+1}^{N}(y_n + \hat{a}_j y_{n-j})^2$$

and the estimates are obtained under $\mathcal{H}_0$ or assuming the signal to be absent.

Figure 1 shows the block diagram of the Rao detector. In the Gaussian case the least squares estimators are known to be close to the MLE and hence any of the least squares techniques (such as the autocorrelation method, the covariance method, the

Forward/backward method and so on) [Kay 1986] can be used to produce the estimators in (48). The numerator appears to be a prewhitener-correlator [Van Trees 1968] except for the squaring required to make the test two-sided. The denominator serves to normalize the statistic so that a constant false alarm rate (CFAR) is maintained asymptotically. Note that it is not necessary to estimate μ or to assume a prior value for it. Also, assuming $a_1, a_2, \cdots, a_p = 0$ or white noise will result in the square a simple correlator or a matched filter in the numerator of ℓ_R, which is known to be the *uniformly most powerful* (UMP) test for the detection of a known signal in white noise [Van Trees 1968]. In the case of the general linear model the moving average (MA) filter with transfer function $A(Z)$ defined by

$$A(Z) = 1 + \sum_{j=1}^{p} a_j Z^{-j}$$

will have to be replaced by an inverse transformation $\mathbf{W}^{-1}$.

VIB. The Mixed-Gaussian Case

The mixed-Gaussian PDF is particularly useful in representing a special class of non-Gaussian noise processes, namely, a nominally Gaussian background contaminated with occasional *impulses* [Sengupta and Kay 1986, 1]. The PDF is given by

$$f(u_n; \sigma_B^2, \sigma_I^2, \epsilon) = \frac{1-\epsilon}{\sqrt{2\pi\sigma_B^2}} e^{\left(-\frac{u_n^2}{2\sigma_B^2}\right)} + \frac{\epsilon}{\sqrt{2\pi\sigma_I^2}} e^{\left(-\frac{u_n^2}{2\sigma_I^2}\right)}$$

where ϵ is a mixture parameter and $0 < \epsilon < 1$. The subscripts B and I are used to denote background and interference, respectively. An alternative formulation of the PDF is

$$f(u_n; \sigma_B^2, \rho, \epsilon) = \frac{1-\epsilon}{\sqrt{2\pi\sigma_B^2}} e^{\left(-\frac{u_n^2}{2\sigma_B^2}\right)} + \frac{\epsilon}{\sqrt{2\pi\rho\sigma_B^2}} e^{\left(-\frac{u_n^2}{2\rho\sigma_B^2}\right)} \tag{49}$$

where $\rho = \sigma_I^2/\sigma_B^2$. Assuming $\rho >> 1$, one can allow for a wide range of amplitudes and frequencies of occurence of spikes by appropriately choosing ρ and ϵ. The Rao detector proposed in this paper is suitable for the mixed-Gaussian PDF (since it is symmetric) and therefore is applicable to many sonar and radar detection problems where the noise background contains impulses resulting from reverberation and clutter, respectively [Kay 1983]. The Rao detector for mixed-Gaussian PDF is now derived. Once again only the AR noise case is considered. It is assumed that σ_B^2 and ρ are known, so that $\Phi = \epsilon$, a scalar. From (49) it follows that

$$\begin{aligned}\frac{f'(u_n;\epsilon)}{f(u_n;\epsilon)} &= -\frac{u_n}{\sigma_B^2}\left[\frac{\frac{1-\epsilon}{\sqrt{2\pi\sigma_B^2}}e^{\left(-\frac{u_n^2}{2\sigma_B^2}\right)}+\frac{\epsilon}{\rho\sqrt{2\pi\rho\sigma_B^2}}e^{\left(-\frac{u_n^2}{2\rho\sigma_B^2}\right)}}{\frac{1-\epsilon}{\sqrt{2\pi\sigma_B^2}}e^{\left(-\frac{u_n^2}{2\sigma_B^2}\right)}+\frac{\epsilon}{\sqrt{2\pi\rho\sigma_B^2}}e^{\left(-\frac{u_n^2}{2\rho\sigma_B^2}\right)}}\right]\\ &= -\frac{u_n}{\sigma_B^2}\Gamma(u_n;\epsilon)\end{aligned} \tag{50}$$

where $\Gamma(u_n;\epsilon)$ represents the term in the brackets. Note that (50) differs from (47) only by the weighting function $\Gamma(u_n;\epsilon)$ which suppresses the high values of u_n [Kay and Sengupta 1986, 2]. In the Gaussian case $\epsilon = 0$ and $\Gamma(u_n;0) = 1$ for *all* u_n. The weighting function $\Gamma(u_n;\epsilon)$, which is an even function of u_n, thus accounts for the non-Gaussian nature of the PDF. In the mixed-Gaussian case, ℓ_R can be calculated by substituting (50) into (35)

$$\ell_R = \frac{\left[\sum_{n=p+1}^{N}\left(s_n+\sum_{j=1}^{p}\hat{a}_j s_{n-j}\right)\left(y_n+\sum_{j=1}^{p}\hat{a}_j y_{n-j}\right)\Gamma\left(y_n+\sum_{j=1}^{p}\hat{a}_j y_{n-j};\hat{\epsilon}\right)\right]^2}{\sigma_B^4 I_f(\hat{\epsilon})\sum_{n=p+1}^{N}\left(s_n+\sum_{j=1}^{p}\hat{a}_j s_{n-j}\right)^2} \tag{51}$$

Note that Γ and I_f depend on the value of $\hat{\epsilon}$. Figure 8.2 shows the block diagram corresponding to (51) and has two apparent differences with Figure 1 (the Gaussian

detector). These are the blocks used to compute Γ and I_f. Substitution of σ_B^2, Γ and I_f by $\hat{\sigma}^2$, 1 and $1/\hat{\sigma}^2$, respectively, would lead one back to the Gaussian case described by (48) and Figure 1. The function Γ can be replaced by a suitable approximation, such as the *Butterworth* function to avoid computation of the exponentials. The MLE's can also be replaced by a reasonably accurate estimator, (*e.g.*, a weighted least squares estimator) [Kay and Sengupta 1986, 2]. With all these simplifications (51) would produce a detector which will reduce computation substantially without a significant loss of performance.

VII. Computer Simulations of Performance of the Rao Detector

This section reports the results of computer simulations of the performance of the Rao detector. The clairvoyant Rao detector is considered to be one basis of comparison, while the theoretical or asymptotic performance is regarded as another. Two AR noise processes are selected for computer simulations. The corresponding parameters are listed in Table A. Process I is broadband while process II is narrowband. A mixed-Gaussian process with $\sigma_B^2 = 1$, $\rho = 100$ and $\epsilon = 0.1$ is chosen as the driving noise. ϵ is assumed to be unknown and so estimated, while σ_B^2 and ρ are assumed to be known. The known part of the signal, *i.e.*, s_n, $n = 1, 2, \cdots, N$ is assumed to be unity. μ is adjusted to yield different values of SNR in the following way. The theoretical noise variance is, as obtained from (49),

$$\sigma^2 = \sigma_B^2[(1-\epsilon) + \epsilon\rho] \tag{52}$$

The noise power P_n of the AR process can be obtained from the stepdown procedure using σ^2 and the process parameters [Kay 1986]. The signal power (actually the signal energy) is the defined as

$$P_s = N\mu^2$$

Defining SNR to be the ratio of P_s and P_n (which represents the SNR at the output

of a correlator), it follows that

$$\mu = \sqrt{\frac{(SNR)P_n}{N}} \tag{53}$$

Thus μ is calculated for a given process such as to produce a desired SNR. N, the number of data points is chosen to be 1000. A probability of false alarm $P_{FA} = 0.01$ is used to evaluate the detection performance. The value of γ' necessary for this is 6.635, as obtained by a search routine so as to satisfy $P_{FA} = 0.01$.

In order to evaluate the asymptotic performance of the Rao test, the noncentrality parameter $\lambda = \lambda_c$ as given by (44) is calculated from μ for each value of SNR as per (53). For the chosen values of the PDF parameters, $\sigma^2 I_f$ is calculated by numerical integration and is found to be 9.0. P_D as defined by (38b) is computed from a table using the values of λ and γ' obtained previously.

The *theoretical* value of the threshold γ', as described above, is used for the Rao detector. This requires one to verify of the theoretical predictions of the asymptotic statistics of ℓ_R under $\mathcal{H}_0$ in order to use the theoretical threshold. Computer simulation results based on 1000 experiments with 1000 data points each using the theoretical threshold in the absence of the signal result in a false alarm in 11 cases for process I and 10 cases for process II, corresponding to an experimental false alarm rate of 0.011 and 0.01, respectively, which are very close to the true value of 0.01. This is expected, since the MLE's when $\mathcal{H}_0$ is true are expected to be more accurate than the MLE's when $\mathcal{H}_1$ is true (since the MLE's are computed under the assumption that $\mathcal{H}_0$ is true). Therefore the use of a theoretical threshold is justified and convenient. The statistic ℓ_R is computed from (51) for 500 different blocks of data, each of length $N = 1000$, for a given SNR. The number of times the statistic exceeds γ', scaled by 500 (the number of experiments), is regarded as the experimental value of the probability of detection. This is repeated for different values of SNR in a suitable range (so as to observe the transition from $P_D = 0$ to $P_D = 1$). The MLE of the AR filter parameters

involved in (51) under $\mathcal{H}_0$ are replaced by the two-stage *weighted least squares* estimator proposed by the authors [Kay and Sengupta 1986, 2] to reduce computation and avoid convergence problems for short data records. Once these parameters are estimated, $\hat{\epsilon}$ is computed as

$$\hat{\epsilon} = \frac{1}{\rho - 1}\left(\frac{\hat{\sigma}^2}{\sigma_B^2} - 1\right)$$

from the prediction error power $\hat{\sigma}^2$ of an MA filter fed by the observed data with coefficients $\hat{a}_1, \hat{a}_2, \cdots \hat{a}_p$ (see (52)). $I_f(\hat{\epsilon})$ is computed for each $\hat{\epsilon}$ from a stored table by interpolation. Γ in the numerator of ℓ_R is replaced by the Butterworth approximation

$$\tilde{\Gamma}(u) = \frac{K_1}{1 + \left|\dfrac{u}{u_c}\right|^{\beta}} + K_2$$

in order to reduce computation. The parameters K_1, K_2, u_c and β are selected to approximate Γ over the range of ϵ.

The statistic ℓ_{RC} for the clairvoyant Rao test, as defined by (40), is computed in a similar way using true values of *all* the parameters. This is used to determine the probability of detection (P_D) of the clairvoyant Rao detector for each SNR via computer simulations.

Figure 3 plots the probabilities of detection of the Rao detector and the clairvoyant Rao detector along with the theoretical or asymptotic performance vs. SNR for the noise process I. Figure 4 plots the same for the noise process II. The three curves exhibit remarkable similarity for each process. It can be concluded that (15) and (38) adequately represent the performance of the computationally efficient detector proposed in this paper, all the approximations incorporated. even for moderately sized data records.

The performance of the Rao detector is now compared to that of a *Gaussian* Rao detector. A Gaussian Rao detector (defined by (48)) is computationally much simpler. The purpose of comparing these two detectors is to show how much one can lose as a

result of an incorrect Gaussian assumption about the driving noise. All least square estimators are close approximations to the MLE if the PDF is assumed to be Gaussian. The covariance method is used to obtain the estimates involved in (48). The same theoretical threshold γ' is used. (Computer simulations based on 1000 experiments with 1000 data points each using the theoretical threshold indicate for process I an experimental false alarm rate of 0.009 which is close to the true value of P_{FA}). The experimental threshold for process II is found to be 0.012 P_D is calculated from the number of times (out of 500 independent trials) ℓ_R exceeds γ'. $N = 1000$ is used and P_D is computed for different values of SNR in a suitable range. Figures 5 and 6 plot the resultant experimental performance of the Gaussian Rao detector in the same scale with those of the mixed-Gaussian Rao detectors, actual and asymptotic, for the two noise processes described before. They indicate a substantial degradation in performance as a result of the Gaussian assumption of the driving noise. It is also observed that the performance of the Gaussian Rao detector matches the predicted asymptotic performance of a Gaussian Rao detector in *Gaussian* noise having equivalent variance. This can be explained in the following way. Since it is assumed that $s_n = 1,\ n = 1, 2, \cdots, N$ (48) can be written as

$$\ell_R = \frac{1}{(N-p)\hat{\sigma}^2} \left[\sum_{n=p+1}^{N} \left(y_n + \sum_{j=1}^{p} a_j y_{n-j} \right) \right]^2 \tag{54}$$

Assuming the estimates of the AR filter parameters to be reasonably accurate, the term in the parantheses is approximately the nth sample of the driving noise which is assumed in this case to be mixed-Gaussian. By Central Limit Theorem arguments the sum of $N-p$ such random variables scaled by $(N-p)\hat{\sigma}^2$ has a Gaussian PDF with variance one assuming $\hat{\sigma}^2$ to be close to the variance of the said mixed-Gaussian PDF. Therefore the asymptotic statistics of ℓ_R in this case will be the same as what it would have been if the noise were Gaussian with the same variance. It can be concluded that the asymptotic statistics of the *Gaussian* Rao detector in *mixed-Gaussian* noise is

represented by (15) and (44), despite modeling error. However this conclusion holds for large data records and small D.C. level signal amplitudes only. A large value of μ is expected to cause a significant degradation of performance of the LS estimators so that the above interpretation of (54) is no longer valid. A short data record will make the Central Limit Theorem inapplicable. Comparing the asymptotic performances of the Gaussian and non-Gaussian detectors, $\sigma^2 \mathcal{I}_f$ is found to describe quantitavely the improvement of the mixed-Gaussian detector over the Gaussian detector when the true noise actually fits the mixed-Gaussian model and the signal is a D.C. level. Figures 5 and 6 show a constant difference of approximately 10 dB between the performances of the Gaussian and mixed-Gaussian detectors which matches the theoretical prediction of 9.6 dB (since $\sigma^2 \mathcal{I}_f \approx 9$ in this case).

VIII. Summary

The Rao efficient score test proposed in this paper is found to be well suited for the problem of detecting a weak signal of unknown amplitude in the presence of colored non-Gaussian noise of unknown PDF and PSD parameters. Since the Rao test is asymptotically equivalent to the GLRT, it shares all the attractive asymptotic properties possessed by the GLRT. It greatly reduces computation by completely avoiding estimation of the unknown parameters under $\mathcal{H}_1$. The Rao detectors were derived for Gaussian and mixed-Gaussian background noise processes. The performance of the Rao detector is found to be equivalent to that of a clairvoyant Rao detector built with perfect knowledge of the nuisance parameters. The experimental performance matches the theoretical predictions of the asymptotic performance for large data records. A detector which assumes the noise PDF to be Gaussian, while it is actually not so, is found to be much inferior in performance.

References

[1] H.L. Van Trees, *Detection, Estimation and Modulation Theory*, Chapter 4, New York: John Wiley, 1968.

[2] D. Sengupta and S.M. Kay, "Efficient Estimation of Parameters for Non-Gaussian Autoregressive Processes", submitted for review to IEEE Trans. on Acoustics, Speech and Signal Processing, 1986.

[3] Sir S.M. Kendall and A. Stuart, *The Advanced Theory of Statistics Vol. II*, Chapters 18–19, New York: Macmillan Publishing, 1979.

[4] S.M. Kay and D. Sengupta, "Simple and Efficient Estimation of Parameters for Non-Gaussian Autoregressive Processes", submitted for review to IEEE Trans. on Acoustics, Speech and Signal Processing, 1986.

[5] C.R. Rao, "Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Applications to Problems of Estimation", Proc. Cam. Phil. Soc., pp. 50–57, Vol. 44, 1948.

[6] A. Wald, "Tests of Statistical Hypotheses Concerning Several Parameters when the Number of Observations is Large", Trans. Am. Math. Soc., pp. 426–482, Vol. 54, 1943.

[7] S.M. Kay and D. Sengupta, "Detection in Incompletely Characterized Colored Non-Gaussian Noise via Parametric Modeling", submitted for review to IEEE Trans. on Acoustics, Speech and Signal Processing, 1986.

[8] C.R. Rao, *Linear Statistical Inference and its Applications*, Chapters 5–6, New York: John Wiley, 1973.

[9] D. Sengupta, "Estimation and Detection for Non-Gaussian Processes using Autoregressive and Other Models", M.S. Thesis, Dept. of Electrical Engineering, Univ. of Rhode Island, 1986.

[10] S.V. Czarnecki and J.B. Thomas, "Nearly Optimal Detection of Signals in Non-Gaussian Noise", ONR report #14, Feb. 1984.

[11] D. Middleton, "Canonically Optimum Threshold Detection", IEEE Trans. on Info. Theory, pp. 230–243, Vol. IT-12, No. 2, April 1966.

[12] D.E. Bowyer *et al*, "Adaptive Clutter Filtering using Autoregressive Spectral Estimation", IEEE Trans. on Aerosp. Electron. Syst., pp. 538–546, July 1979.

[13] S.M. Kay, "Asymptotically Optimal Detection in Unknown Colored Noise via Autoregressive Modeling", IEEE Trans. on Acoustics, Speech and Signal Processing, pp. 927–940, Vol. ASSP-31, Aug. 1983.

[14] S.M. Kay, "Asymptotically Optimal Detection in Incompletely Characterized Non-Gaussian Noise", Submitted for review to IEEE Trans. on Acoustics, Speech and Signal Processing, 1985.

[15] S.M. Kay, *Modern Spectral Estimation: Theory and Application*, Chapter 6, Prentice-Hall, to be published (1986).

[16] D.R. Cox and D.V. Hinkley, *Theoretical Statistics*, Chapter 9, London: Chapman and Hall, 1974.

Table A: Parameters of the AR processes used for simulation

Process	a_1	a_2	a_3	a_4	poles
I	-1.352	1.338	-0.662	0.240	$0.7\exp[j2\pi(0.12)]$ $0.7\exp[j2\pi(0.21)]$
II	-2.760	3.809	-2.654	0.924	$0.98\exp[j2\pi(0.11)]$ $0.98\exp[j2\pi(0.14)]$

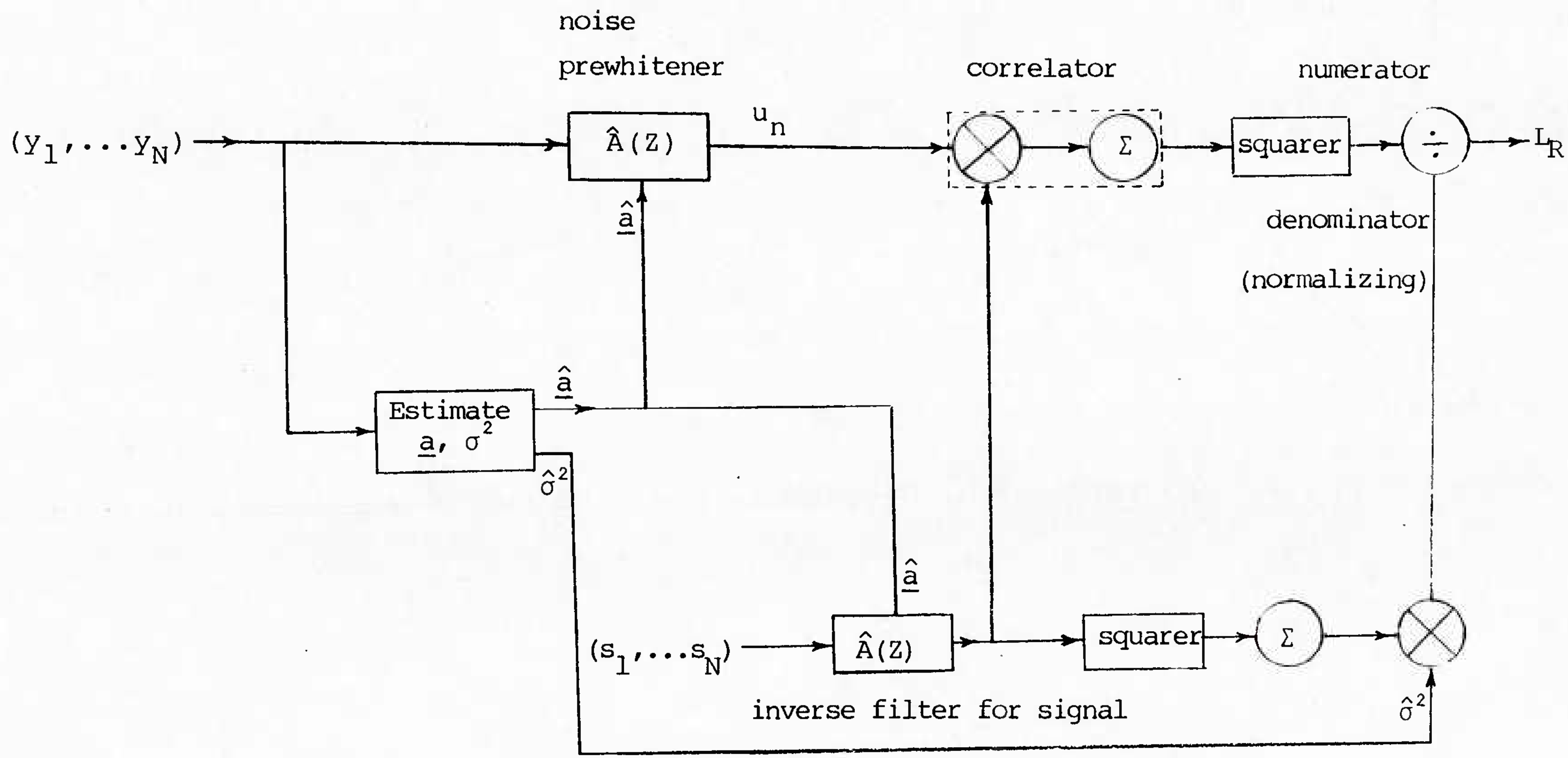

Figure 1 Block diagram of Rao detector for Gaussian noise

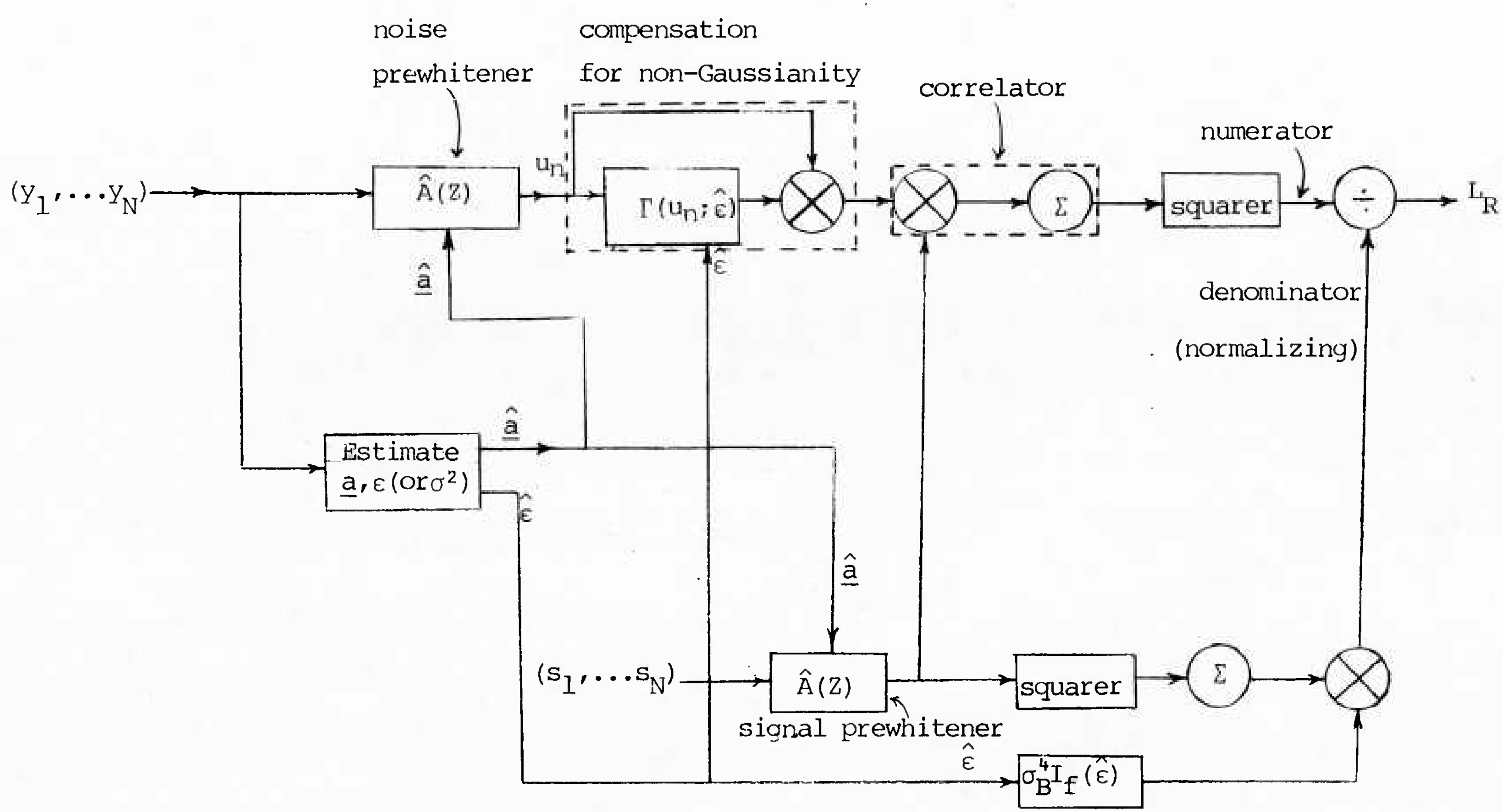

Figure 2 Block diagram of Rao detector for mixed-Gaussian noise

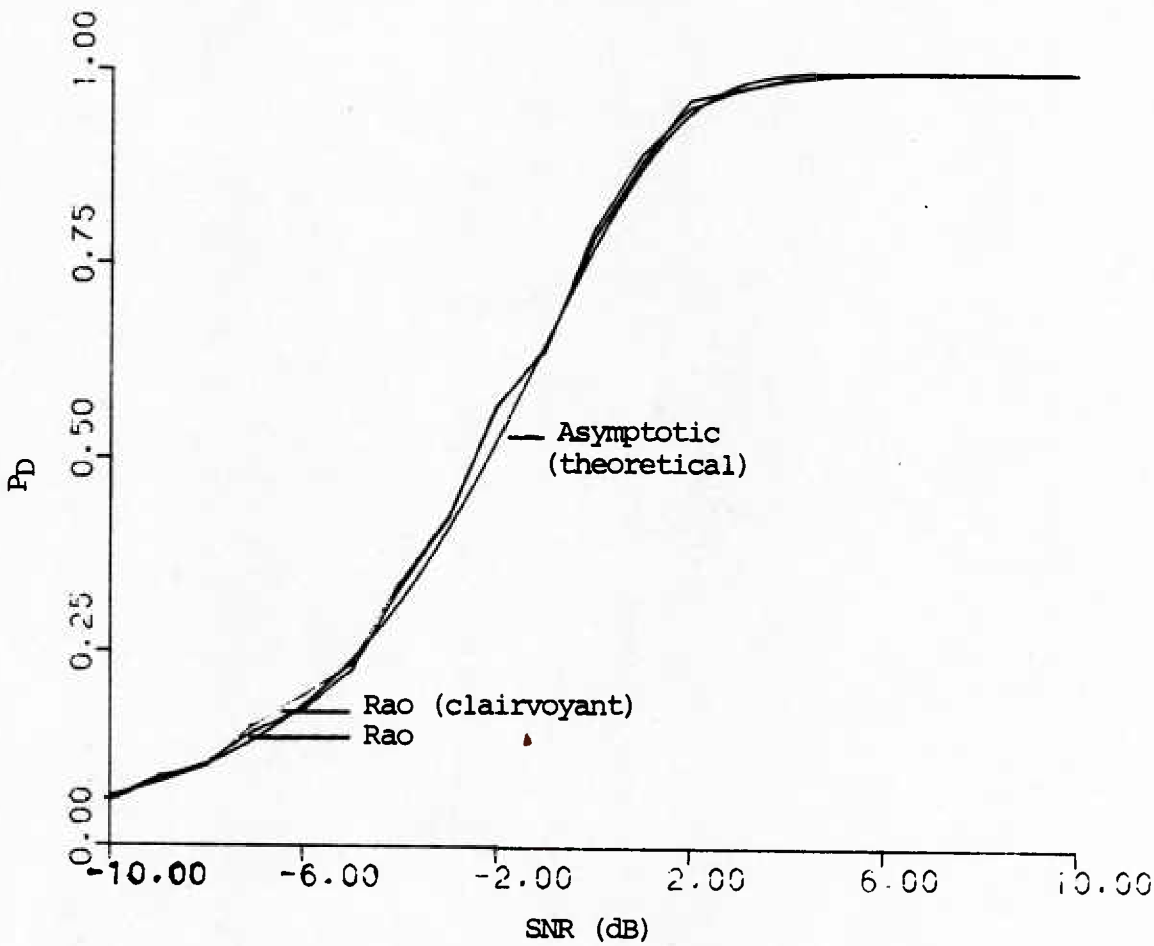

Figure 3 Performance of the Rao detector for noise process I

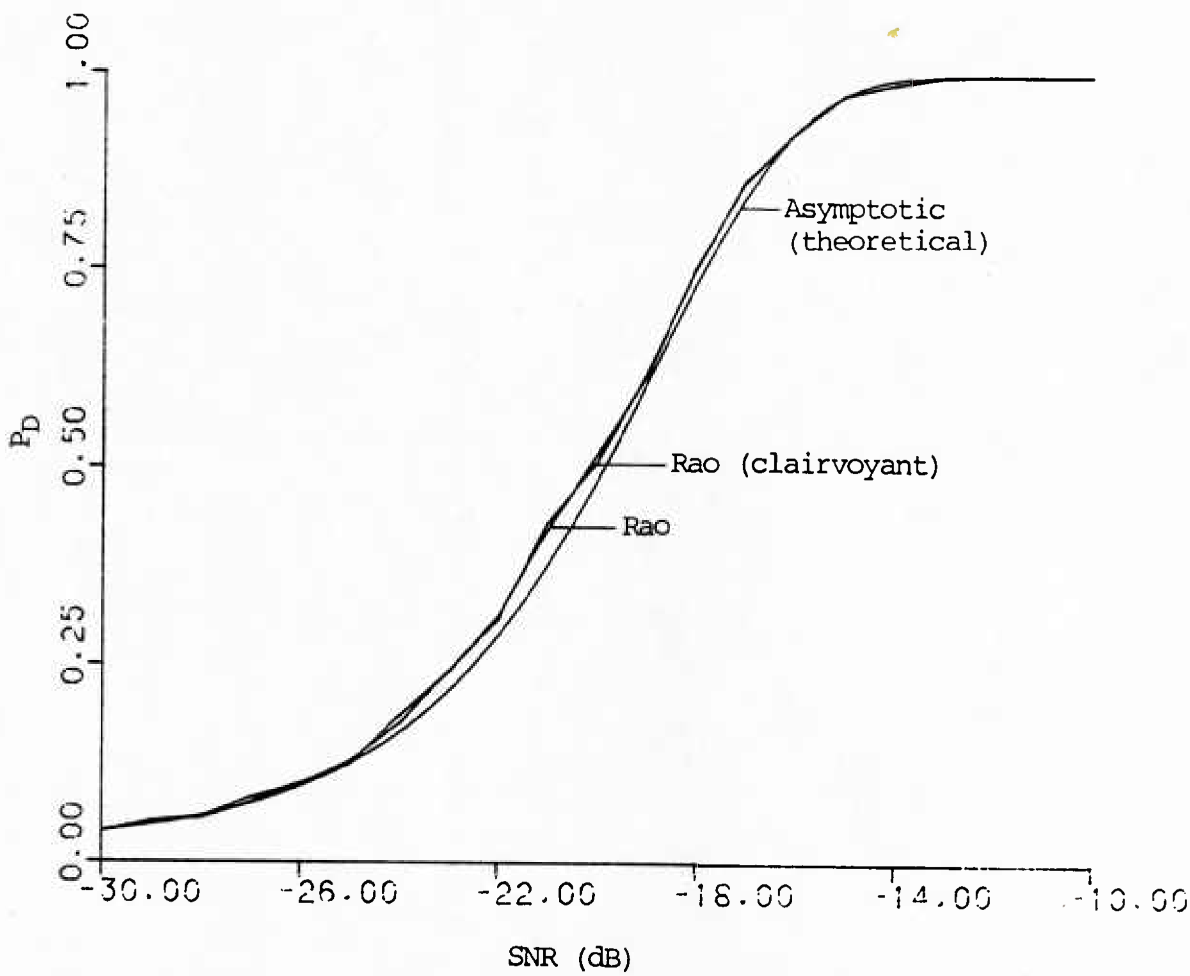

Figure 4 Performance of the Rao detector for noise process II

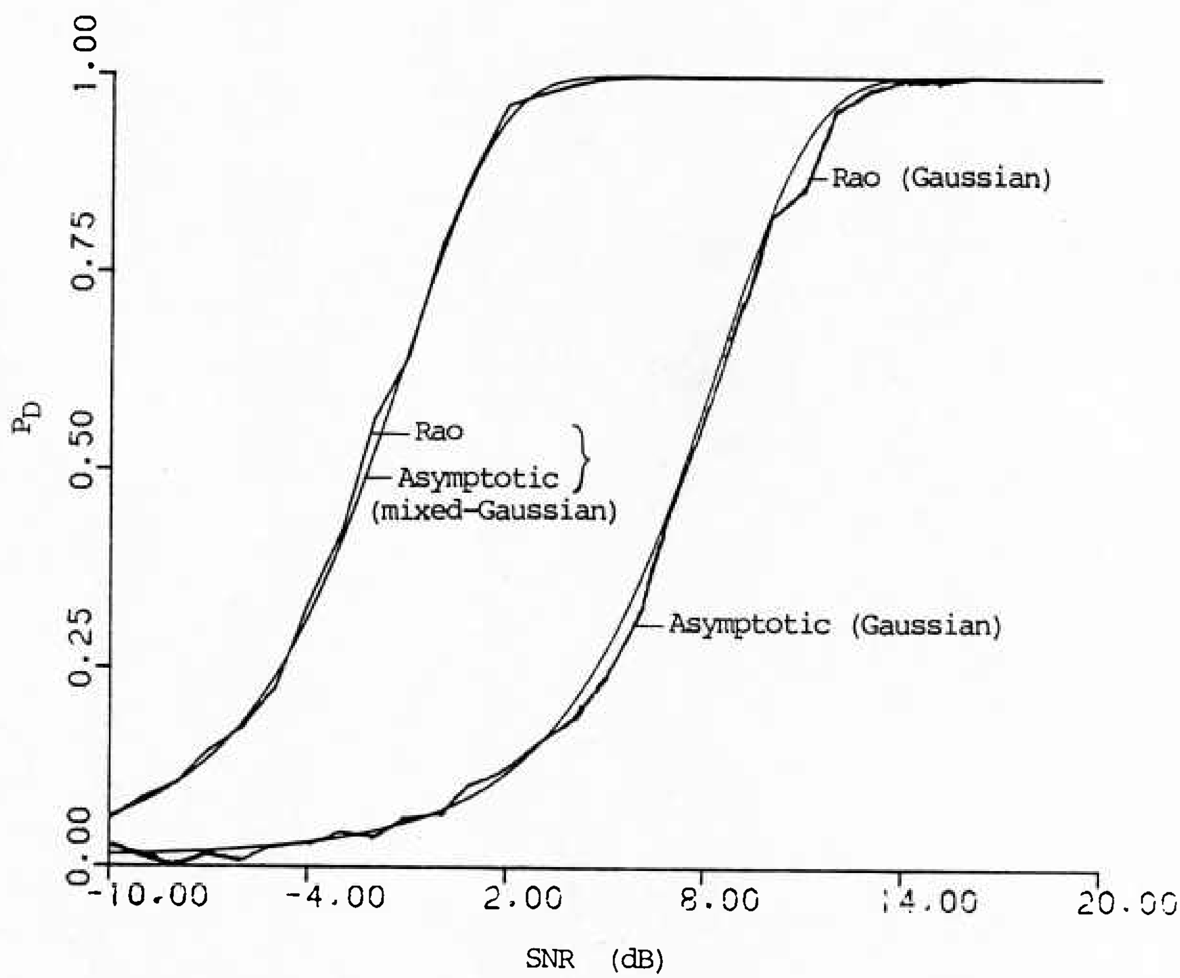

Figure 5 Performance of the Gaussian and mixed-Gaussian Rao detectors for mixed Gaussian noise process I

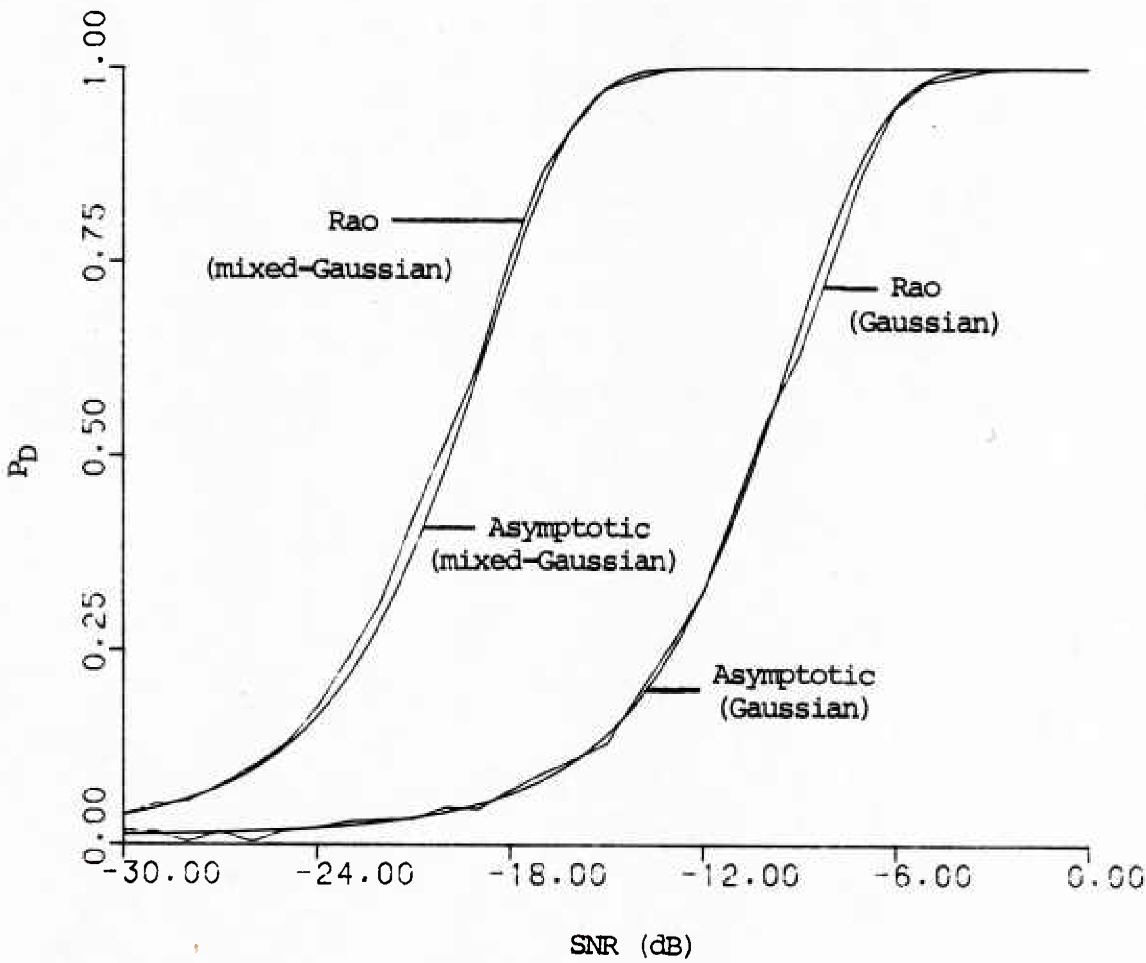

Figure 6 Performance of the Gaussian and mixed-Gaussian Rao detectors for mixed-Gaussian noise process II

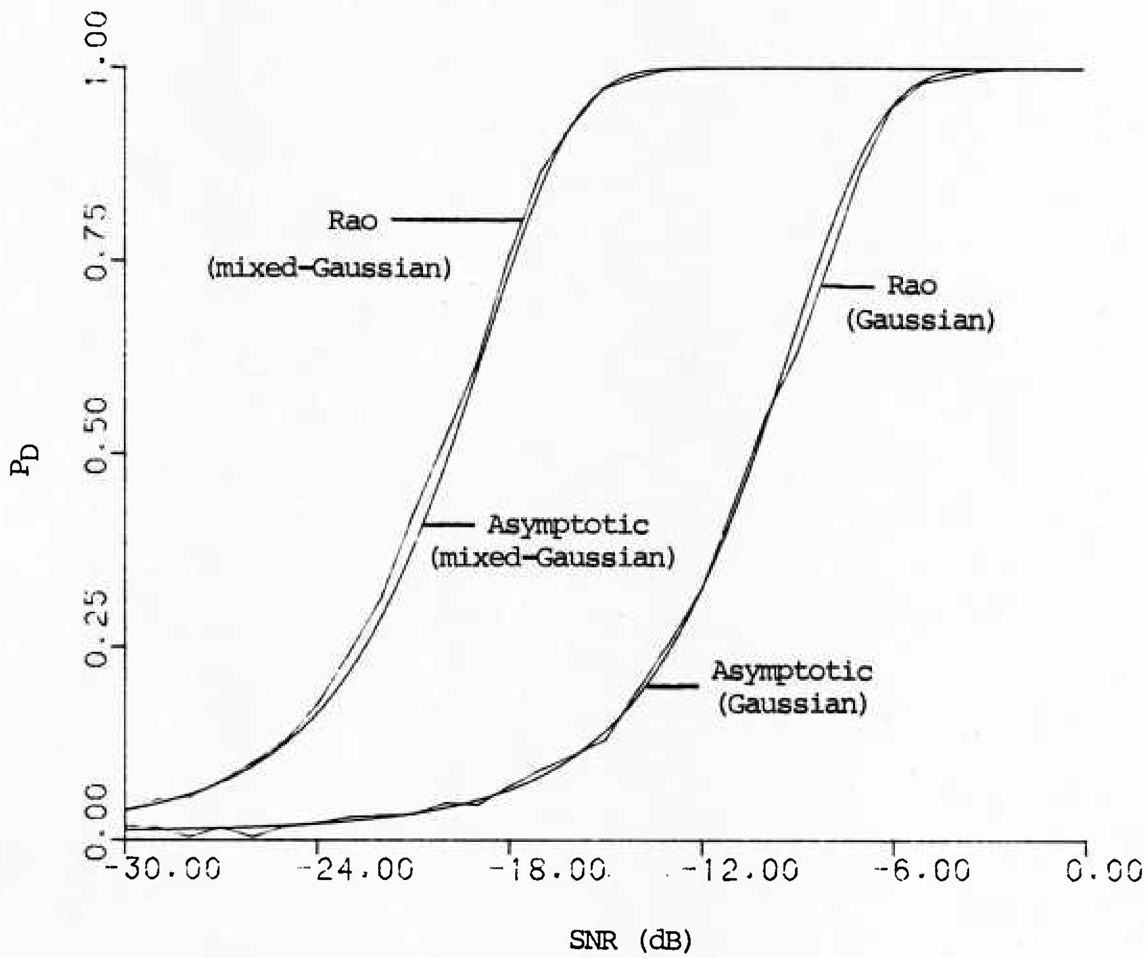

Figure 6 Performance of the Gaussian and mixed-Gaussian Rao detectors for mixed-Gaussian noise process II

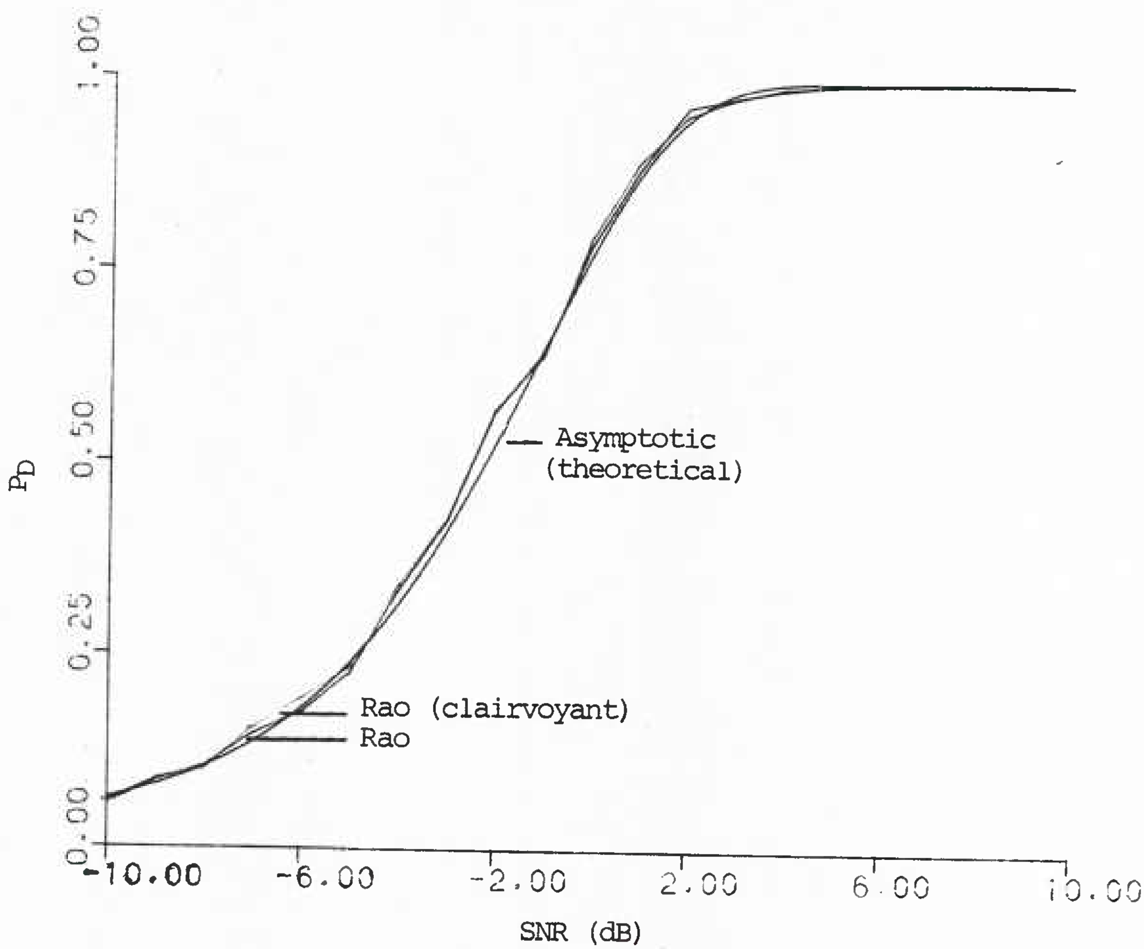

Figure 3 Performance of the Rao detector for noise process I

OFFICE OF NAVAL RESEARCH
STATISTICS AND PROBABILITY PROGRAM

BASIC DISTRIBUTION LIST
FOR
UNCLASSIFIED TECHNICAL REPORTS

FEBRUARY 1982

	Copies
Statistics and Probability Program (Code 411(SP)) Office of Naval Research Arlington, VA 22217	3
Defense Technical Information Center Cameron Station Alexandria, VA 22314	12
Commanding Officer Office of Naval Research Eastern/Central Regional Office Attn: Director for Science Barnes Building 495 Summer Street Boston, MA 02210	1
Commanding Officer Office of Naval Research Western Regional Office Attn: Dr. Richard Lau 1030 East Green Street Pasadena, CA 91101	1
U. S. ONR Liaison Office - Far East Attn: Scientific Director APO San Francisco 96503	1
Applied Mathematics Laboratory David Taylor Naval Ship Research and Development Center Attn: Mr. G. H. Gleissner Bethesda, Maryland 20084	1
Commandant of the Marine Coprs (Code AX) Attn: Dr. A. L. Slafkosky Scientific Advisor Washington, DC 20380	1
Navy Library National Space Technology Laboratory Attn: Navy Librarian Bay St. Louis, MS 39522	1
U. S. Army Research Office P.O. Box 12211 Attn: Dr. J. Chandra Research Triangle Park, NC 27706	1
Director National Security Agency Attn: R51, Dr. Maar Fort Meade, MD 20755	1
ATAA-SL, Library U.S. Army TRADOC Systems Analysis Activity Department of the Army White Sands Missile Range, NM 88002	1
ARI Field Unit-USAREUR Attn: Library c/o ODCSPER HQ USAEREUR & 7th Army APO New York 09403	1
Library, Code 1424 Naval Postgraduate School Monterey, CA 93940	1
Technical Information Division Naval Research Laboratory Washington, DC 20375	1
OASD (I&L), Pentagon Attn: Mr. Charles S. Smith Washington, DC 20301	1

Copies

Director
AMSAA
Attn: DRXSY-MP, H. Cohen
Aberdeen Proving Ground, MD 21005 — 1

Dr. Gerhard Heiche
Naval Air Systems Command (NAIR 03)
Jefferson Plaza No. 1
Arlington, VA 20360 — 1

Dr. Barbara Bailar
Associate Director, Statistical Standards
Bureau of Census
Washington, DC 20233 — 1

Leon Slavin
Naval Sea Systems Command (NSEA 05H)
Crystal Mall #4, Rm. 129
Washington, DC 20036 — 1

B. E. Clark
RR #2, Box 647-B
Graham, NC 27253 — 1

Naval Underwater Systems Center
Attn: Dr. Derrill J. Bordelon
Code 601
Newport, Rhode Island 02840 — 1

Naval Coastal Systems Center
Code 741
Attn: Mr. C. M. Bennett
Panama City, FL 32401 — 1

Naval Electronic Systems Command (NELEX 612)
Attn: John Schuster
National Center No. 1
Arlington, VA 20360 — 1

Defense Logistics Studies Information Exchange
Army Logistics Management Center
Attn: Mr. J. Dowling
Fort Lee, VA 23801 — 1

Copies

Reliability Analysis Center (RAC)
RADC/RBRAC
Attn: I. L. Krulac
Data Coordinator/
Government Programs
Griffiss AFB, New York 13441 — 1

Technical Library
Naval Ordnance Station
Indian Head, MD 20640 — 1

Library
Naval Ocean Systems Center
San Diego, CA 92152 — 1

Technical Library
Bureau of Naval Personnel
Department of the Navy
Washington, DC 20370 — 1

Mr. Dan Leonard
Code 8105
Naval Ocean Systems Center
San Diego, CA 92152 — 1

Dr. Alan F. Petty
Code 7930
Naval Research Laboratory
Washington, DC 20375 — 1

Dr. M. J. Fischer
Defense Communications Agency
Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090 — 1

Mr. Jim Gates
Code 9211
Fleet Material Support Office
U. S. Navy Supply Center
Mechanicsburg, PA 17055 — 1

Mr. Ted Tupper
Code M-311C
Military Sealift Command
Department of the Navy
Washington, DC 20390 — 1